

ДЕТЕКТИРОВАНИЕ ОБЪЕКТОВ НА ИЗОБРАЖЕНИИ НА ОСНОВЕ КОМБИНАЦИИ HOG+SVM

В.М. Рябов, Ю.А. Иванова

Томский политехнический университет

E-mail: viktor.ryabov.m@gmail.com, jbolotova@tpu.ru

Аннотация

В работе представляется алгоритм классификации изображений на основе комбинации гистограмм направленных градиентов (HOG) и машины опорных векторов (SVM). Для оценки качества работы алгоритма применяется процентная точность. Проведены сравнения разных запусков обучения алгоритма с использованием разных гиперпараметров.

Введение

Алгоритмы обнаружения объектов, основанные на компьютерном зрении, обычно состоят из детектора признаков изображения, который обеспечивает цифровое описание, и распознавателя образов, который классифицирует изображения. Алгоритмы обнаружения объектов могут применяться для различных диапазонов объектов, ориентаций и условий освещения.

Существует множество алгоритмов, используемых для этих целей. Некоторые из них весьма плодотворны в конкретных областях.

В этой работе была выбрана тематика классификации людей на изображениях с использованием комбинации методов HOG+SVM [1].

Идея HOG состоит в том, чтобы вместо использования каждого отдельного направления градиента каждого отдельного пикселя изображения, группировать пиксели в маленькие ячейки. Для каждой ячейки мы вычисляем все направления градиента и группируем их в несколько бинов ориентации. Суммируем величину градиента в каждом образце. Таким образом, более сильные градиенты вносят больший вес в свои бункеры, а эффекты малых случайных ориентаций из-за шума снижаются. Эта гистограмма дает нам картину доминирующей ориентации этой клетки. Выполнение этого для всех ячеек дает нам представление о структуре изображения.

На полученных данных HOG будет обучаться линейный классификатор SVM.

Все программы реализованы на языке Python.

Описание выборки данных

В данной работе используется выборка данных INRIA Person [2]. Выборка разделена на тестовую и тренировочную. В свою очередь они разделены на негативные и позитивные изображения, на которых изображены люди, и на которых люди отсутствуют. При обучении и тестировании сети использовалась выборка из 2 573 изображений

разного размера. В обучающей выборке использовалось 1832 изображения, из которых 612 были позитивные, а 1218 негативными примерами. В тестовой выборке использовались 735 изображений, из которых 288 позитивных, а 453 негативные примеры.

На данных изображениях будут определяться градиенты и классифицировать, есть на изображении человека или его нет.

Пример изображения из тестовой выборки, можно увидеть на изображении, которое находится ниже (рис 1):

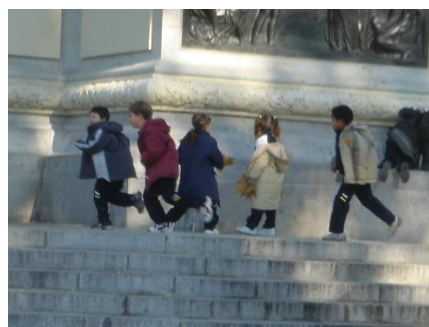


Рис. 1. Пример входного изображения

Реализация алгоритмов

На первом этапе применяется необязательное глобальное выравнивание нормализации изображения, предназначенное для уменьшения влияния эффектов освещения [3]. Второй этап вычисляет градиенты изображения первого порядка. Они фиксируют контур, силуэт и некоторую информацию о текстуре, обеспечивая дополнительную устойчивость к изменениям освещенности. Третий этап направлен на создание кодирования, чувствительного к локальному содержанию изображения, в то же время оставаясь устойчивым к небольшим изменениям позы или внешнего вида. Окно изображения делится на небольшие пространственные области, называемые «ячейками». Четвертый этап вычисляет нормализацию, которая берет локальные группы клеток и, напротив, нормализует их общие ответы перед переходом к следующему этапу. Обычно каждая отдельная ячейка распределяется между несколькими блоками, но ее нормализация зависит от блока и, следовательно, отличается. Таким образом, ячейка появляется несколько раз в конечном выходном векторе с различной нормализацией.

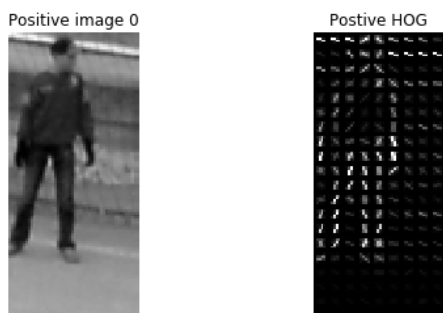


Рис. 2. Результат работы алгоритма HOG

На (рис. 2) мы видим результат работы алгоритма HOG. Мы можем менять параметры orientation, pix_per_cell и cell_per_block для вычисления характеристик HOG одного канала изображения [4]. Количество ориентаций — это количество бинов, градиентов пикселей каждой ячейки, которые будут разделены на гистограмме. Pix_per_cells — это количество пикселей в каждой строке и столбце на ячейку для каждого градиента, который вычисляется гистограммой.

Cell_per_block определяет локальную область, по которой гистограмма рассчитывает в данной ячейки. Также используется схема нормализации под названием transform_sqrt, которая помогает уменьшить влияние теней и изменений освещенности [5].

Результаты работы

Было проведено обучение и тестирование комбинации алгоритмов на основе комбинации HOG+SVM. Точность классификации алгоритма при разных значениях гиперпараметров представлены в таблице ниже (таблица 1):

Таблица 4. Точность классификации алгоритма при разных значениях гиперпараметров

orientations=9 pixels_per_cell=(8, 8) cells_per_block=(2, 2)	65,18%
orientations=24 pixels_per_cell=(12, 12) cells_per_block=(2, 2)	68,15%
orientations=64 pixels_per_cell=(7, 7) cells_per_block=(2, 2)	75,84%

Результат работы алгоритма при правильной классификации. (рис. 3):

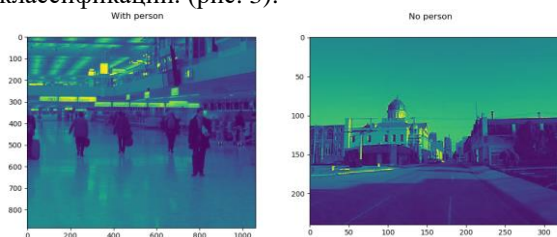


Рис. 3. Результат правильно классификации.

Исходя из полученных результатов, можно заметить, что при повышении параметра Orientation следствием будет увеличение точности, потому как повышается количество градиентов пикселей каждой ячейки. Также параметр pixels_per_cell тоже влияет на точность, так как имеет значение количество пикселей в каждой строке и столбце для каждого градиента.

Заключение

Разработанная программа позволяет классифицировать изображения и при корректном подборе гиперпараметров и обучающей выборки хорошо справляется с данной задачей. Комбинация HOG+SVM является эффективным и простым алгоритмом.

Полученная точность может быть улучшена путем использования большего количества (не избыточных) обучающих примеров, а также лучшей подборки гиперпараметров.

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 18-08-00977 А и в рамках Программы повышения конкурентоспособности ТПУ.

Список использованных источников:

1. Хабр [Электронный ресурс]: Классификация данных методом опорных векторов. URL: <https://habr.com/ru/post/105220/> (21.11.2019)
2. INRIA Person Dataset [Электронный ресурс]. — Режим доступа: URL: <http://pascal.inrialpes.fr/data/human/> (21.11.2019)
3. Navneet Dalal, Bill Triggs— Histograms of Oriented Gradients for Human Detection// INRIA Rhone-Alps/ (21.11.2019)
4. Shah Nawaz - HOG-SVM Car Detection on an Embedded GPU// ResearchGate. (11.12.2019)
5. Siji Joseph , Arun Pradeep - Object Tracking using HOG and SVM// Axis College of Engineering and Technology/(11.12.2019)